

**This is a list of steps required to analyze DNA microarray data.**

1. Quantitate the fluorescence signal in each spot
2. Calculate the ratio of red/green fluorescence
3. Log transform the ratios
4. Normalize the ratios on each microarray slide
5. Normalize the ratios for a set of slides in an experiment
6. Perform statistical analysis on the ratios

Steps 1-6 have already been performed

7. Compare individual genes with known data

Step 7 was performed by the research group who gathered the data

We are performing this step:

8. Map onto biological pathways (GenMAPP & MAPPFinder)

**GenMAPP Expression Dataset Manager Procedure**

1. Launch the GenMAPP Program. Select the *Data* menu from the main Drafting Board window and select *Choose Gene Database* from the drop-down list. Navigate to the location C:\GenMAPP 2 Data\Gene Databases\Hs-Std\_20070817.gdb as your Gene Database. This is the Gene Database for humans.
2. Select the *Data* menu from the main Drafting Board window and choose *Expression Dataset Manager* from the drop-down list. The Expression Dataset Manager window will open.
3. Select New Dataset from the Expression Datasets menu. Select the .txt file Hs\_ED\_ProstateCancer\_edited.txt from the file dialog box that appears. (This file can be downloaded from the Course Wiki under the section "Lecture Slides & Other Handouts").
4. The Data Type Specification window will appear. The MAPP Location field has text (character) data, so check that box. None of the other columns have text data, so leave those boxes unchecked and click OK.
5. Allow the Expression Dataset Manager to convert your data.  
*This may take a few minutes depending on the size of the dataset and the computer's memory and processor speed. When the process is complete, the converted dataset will be active in the Expression Dataset Manager window and the file will be saved in the same folder the raw data file was in, named the same except with a .gex extension; for example, MyExperiment.gex. A message may appear saying that the Expression Dataset Manager could not convert one or more lines of data. Lines that generate an error during the conversion of a raw data file are not added to the Expression Dataset. Instead, an exception file is created. The exception file is given the same name as your raw data file with .EX before the extension (e.g., MyExperiment.EX.txt). The exception file will contain all of your raw data, with the addition of a column named ~Error~. This column contains either error messages or, if the program finds no errors, a single space character. We will not worry about conversion errors at this time.*
6. Customize the new Expression Dataset by creating new Color Sets which contain the instructions to GenMAPP for displaying data on MAPPs.

7. Color Sets contain the instructions to GenMAPP for displaying data from an Expression Dataset on MAPPs. Create a Color Set by filling in the following different fields in the Color Set area of the Expression Dataset Manager: a name for the Color Set, the gene value, and the criteria that determine how a gene object is colored on the MAPP. Enter a name in the Color Set Name field that is 20 characters or fewer.
8. The Gene Value is the data displayed next to the gene box on a MAPP. Select the column of data to be used as the Gene Value from the drop down list or select [none].
9. Activate the Criteria Builder by clicking the New button.
10. Enter a name for the criterion in the Label in Legend field.
11. Choose a color for the criterion by left-clicking on the Color box. Choose a color from the Color window that appears and click OK.
12. State the criterion for color-coding a gene in the Criterion field.  
*A criterion is stated with relationships such as "this column greater than this value" or "that column less than or equal to that value". Individual relationships can be combined using as many ANDs and ORs as needed. A typical relationship is*  
*[ColumnName] RelationalOperator Value*  
*with the column name always enclosed in brackets and character values enclosed in single quotes. For example:*  
*[Fold Change] >= 2*  
*[p value] < 0.05*  
*[Quality] = 'high'*  
*The easiest and safest way to create criteria is by choosing items from the Columns and Ops (operators) lists shown in the Criteria Builder. The Columns list contains all of the column headings from your Expression Dataset. To choose a column from the list, click on the column heading. It will appear at the location of the cursor in the Criterion box. The Criteria Builder surrounds the column names with brackets.*  
*The Ops (operators) list contains the relational operators that may be used in the criteria: equals ( = ) greater than ( > ), less than ( < ), greater than or equal to ( >= ), less than or equal to ( <= ), is not equal to ( <> ). To choose an operator from the list, click on the symbol. It will appear at the location of the insertion bar (cursor) in the Criterion box. The Criteria Builder automatically surrounds the operators with spaces.*  
*The Ops list also contains the conjunctions AND and OR, which may be used to make compound criteria. For example:*  
*[Fold Change] > 1.2 AND [p value] <= 0.05*  
*Parentheses control the order of evaluation. Anything in parentheses is evaluated first. Parentheses may be nested. For example:*  
*[Control Average] = 100 AND ([Exp1 Average] > 100 OR [Exp2 Average] > 100)*  
*Column names may be used anywhere a value can, for example:*  
*[Control Average] < [Experiment Average]*
13. After completing a new criterion, add the criterion entry (label, criterion, and color) to the Criteria List by clicking the Add button.
14. Continue to add criteria to the Color Set by following steps 11-14.

*The buttons to the right of the list represent actions that can be performed on individual criteria. To modify a criterion label, color, or the criterion itself, first select the criterion in the list by left-clicking on it, and then click the Edit button. This puts the selected criterion into the Criteria Builder to be modified. Click the Save button to save changes to the modified criterion; click the Add button to add it to the list as a separate criterion. To remove a criterion from the list, left-click on the criterion to select it, and then click on the Delete button. The order of Criteria in the list has significance to GenMAPP. When applying an Expression Dataset and Color Set to a MAPP, GenMAPP examines the expression data for a particular gene object and applies the color for the first criterion in the list that is true. Therefore, it is imperative that when criteria overlap the user put the most important or least inclusive criteria in the list first. To change the order of the criteria in the list, left-click on the criterion to select it and then click the Move Up or Move Down buttons. No criteria met and Not found are always the last two positions in the list.*

15. Save the entire Expression Dataset by selecting Save from the Expression Dataset menu. Changes made to a Color Set are not saved until you do this.
16. Exit the Expression Dataset Manager to view the Color Sets on a MAPP. Choose Exit from the Expression Dataset menu or click the close box in the upper right hand corner of the window.